G. Mandolino · A. Carboni · S. Forapani V. Faeti · P. Ranalli

Identification of DNA markers linked to the male sex in dioecious hemp (Cannabis sativa L.)

Received: 16 January 1998 / Accepted: 30 April 1998

Abstract A 400-bp RAPD marker generated by a primer of random decamer sequence has been found associated with the male sex phenotype in 14 dioecious cultivars and accessions of hemp (*Cannabis sativa* L.). The primer OPA8 generates a set of bands, most of which polymorphic among all the individual plants tested, and 1 of which, named $OPA8₄₀₀$, present in all male plants and absent in female plants. A screening of 167 plants belonging to different genotypes for the association of the $OPA8₄₀₀$ marker with the sex phenotype revealed that only in 3 cases was the 400-bp band was present in plants phenotypically female; on the contrary, in male plants the band was never missing, while in monoecious plants it was never present. Despite this sex-specific association, the sequences corresponding to $OPA8₄₀₀$ were present in both staminate and carpellate plants, as revealed by Southern blotting and hybridization with the cloned RAPD band. The RAPD marker was sequenced, and specific primers were constructed. These primers generated, on the same genotypes used for RAPD analysis, a SCAR marker 390 bp in length and male-specific. This SCAR is suitable for a precise, early and rapid identification of male plants during breeding programs of dioecious and monoecious hemp.

Key words *Cannabis sativa* · Dioecy · Sex · RAPD · SCAR

Introduction

The genus *Cannabis* is one of the most ancient cultivated crops utilized for fibre production and drug

G. Mandolino (\boxtimes) · A. Carboni · S. Forapani V. Faeti *·* P. Ranalli Istituto Sperimentale per le Colture Industriali, Via di Corticella 133, 40129 Bologna, Italy Fax: #39-051-374857 E-mail: istsci10@iperbole.bologna.it

preparation and more recently for oil extraction from the seeds and for paper manifacturing (Abel 1990). All *Cannabis* species are dioecious in nature, and from a caryological point of view all hemp forms are reported to have a $2n = 20$ chromosome set and are supposed to have male heterogamety, though the latter is still controversial (Yamada 1943; Mohan Ram and Sett 1985).

Until recently, the hemp cultivars employed have ben typically dioecious, but monoecious cultivars have been developed as well. Monoecious hemp is particularly useful when both seed production and stem harvesting is performed; the main disadvantages in the use of monoecious hemp essentially lie in the occurrence of self-pollination, causing a lower stem yield, and in the slower rate of genetic progress attainable compared to dioecious hemp (Bócsa 1994). Besides, only in dioecious hemp can the Bredemann principle be applied, which consists in the scoring of the fibre content in male plants and the subsequent pollination of only those plants with the highest fibre content (Bredemann 1938). As the adoption of this method implies a direct identification of the male plants, which is possible only when reproductive maturity approaches, the risks of contamination resulting in seeds obtained by undesired pollinations is not negligible. For these reasons it would be very useful in hemp genetic improvement programs to be able to achieve an early recognition of male or female plants, consequently anticipating the fibre content analysis well before the flowering time or designing the crosses to be performed early during the vegetative phase of the plants.

In a previous paper (Faeti et al. 1996), an analysis of the genetic variability in different hemp cultivars and accessions was carried out based on randomly amplified polymorphic DNA (RAPD) markers. In that work, 54 hemp plants belonging to 13 cultivars and accessions of different origins were analyzed; out of 205 scorable markers, only 4 were monomorphic among all the cultivars and accessions tested. Within the

Communicated by G. Wenzel

'Carmagnola' and 'Fibranova' cultivars, the level of polymorphic markers decreased to 80%, which is still a very high value considering the genetic relations between these two varieties (Allavena 1961). These results are not unexpected in a dioecious, obligate outbred species that has been scarcely subjected to selection in the last decades since hemp breeding has virtually ceased because of rigid legislation constraints. It was also shown that, at least in the two mentioned cultivars, male and female plants tended to form two distinct clusters. These preliminary results encouraged us to look for specific molecular markers linked to the sex phenotype in dioecious hemp, as already reported by other authors in this species (Sakamoto et al. 1995). The work presented here describes a single DNA fragment generated by one decamer primer that appears to be closely associated to the male phenotype in different hemp cultivars and accessions and its transformation into a sequence-characterized amplified region (SCAR) marker allowing a rapid and specific identification of the male sex.

Materials and methods

Table 1 Hemp germplasm examined by RAPD analysis

Plant material and DNA preparation

Seeds of the hemp varieties and accessions were obtained from the Institutions listed in Table 1, where the geographical origins of the plants are also specified.

Plants were grown in pots in a greenhouse under natural light and photoperiod. Leaf samples were collected (about 0.5 g) from 7-week old individual plants and stored at -75° C. Hemp plants were then transferred to the field until they flowered and then scored for the sex phenotype. Total genomic DNA was prepared from the collected leaves according to Taylor and Powell (1985) with minor modifications (Faeti et al. 1996).

Polymerase chain reaction (PCR) analysis

RAPD analysis were performed with 20 decamer primers on 50 ng of DNA, subjected to 35 cycles of amplification, each cycle consisting of 1 min at 94*°*C, 2 min at 38*°*C (annealing) and 2 min at 72[°]C (extension), in the presence of 1 U of *Taq* polymerase (Boehringer Mannheim, Germany), 2.5 mM MgCl₂ 0.125 mM each dNTPs 50 ng/µl gelatin (Sigma, USA) 2.5 pmoles of 20 different primers of the A and G series (Operon Technologies, USA). To test the stability and reproducibility of the scored RAPD marker, we also tested different annealing temperatures (38, 41 and 44*°*C). SCAR markers were produced using the same protocol, with the exception that the concentration of $MgCl₂$ in the amplification mixture was 1.5 m*M*, and the annealing temperature was 60*°*C.

Cloning and sequencing

The putatively male-specific band amplified by RAPD analysis was excised from the gel, eluted by the Gene Clean procedure (Bio 101, USA), cloned in the pT7 vector (Novagen, USA) and the chimaeric plasmids inserted in *E*. *coli* cells (strain XL1 Blue, Stratagene, USA). The inserts of the positive clones were amplified by PCR using the primers T7 and U19 flanking the polylinker region of the plasmid vector. A screening of the positive clones showed that three classes of inserts were amplified, each consisting of a single DNA fragment with an approximate size between 400 and 450 bp. The resulting fragments were phenol-purified, precipitated and used as probes in the Southern blot experiments. The same DNA fragments in the cloned form were sequenced by the dideoxynucleotide chain termination method using T7 and U19 as forward and reverse primers. The University of Wisconsin Genetics Computer Group program library (Madison, Wis., USA) was used for nucleic acid sequence comparison; the database search was made by using the BLAST algorithm (Altschul et al. 1990).

Southern blotting

For genomic Southern blot analysis, 10 µg of genomic DNA from 17 male and 28 female plants (chosen among those exhibiting

!ISCI, Istituto Sperimentale per le Colture Industriali, Bologna, Italy; GATE, Rudolf Fleischmann Agricultural Research Center, Kompolti, Hungary; INF, Institute of Natural Fibres, Póznan, Poland; IPK, Institut für Pflanzengenetik und Kulturpflanzenforschung, Gatersleben, Germany

polymorphism for the RAPD marker) were digested to completion with the enzymes *Hin*dIII, *Eco*RV, *Bam*HI, *Xba*I, *Pst*I, *Nco*I, *Bgl*II, *Hin*dII, *Xho*I, *Cla*I, *Pvu*II, *Msp*I and *Hae*III (20 U) for 6 h at 37*°*C. The digests were then ethanol-precipitated, resuspended in distilled water and run on a 1% TAE agarose gel at 2.5 V/cm. The gels were then depurinated, denatured and neutralized according to Sambrook et al. (1989) and blotted by capillary action onto nylon membranes (Boehringer Mannheim, Germany). To check the identity of the cloned fragments with the RAPD band, we also prepared Southern blots of the RAPD amplification products by the same procedure. The cloned fragments of the three different classes were PCR-amplified as described above, ethanol-precipitated and labeled with α -[³²P]dCTP according to Feinberg and Vogelstein (1983). The hybridization protocol was according to Sambrook et al. (1989); after the last wash, the filters were exposed on Kodak autoradiography films type X-Omat AR5.

Results

RAPD marker

From a preliminary study, it was found that the decamer primer OPA8 (5'-GTGACGTAGG-3'), when tested on the dioecious cultivars 'Carmagnola' and 'Fibranova', produced a band of approximately 400 bp that appeared to be male-specific. To check the possibility that the OPA8400 band was linked to the male phenotype, we screened a group of 167 plants belonging to seven commercial cultivars (5 dioecious and 2 monoecious) and seven accessions from germplasm banks or collections (Table 1) for the presence or absence of this 400-bp marker. The results are summarized in Table 2. Figure 1 shows the amplification patterns produced by the primer OPA8 on 30 of the plants tested. The presence of the 400-bp band is clearly evident in all male plants, as is its absence in female plants. Only in 2 cases out of 36 in cv 'Carmagnola', and in 1 case out of 22 in cv 'Fibranova' did a female plant show the OPA8400 marker (Fig. 1b, lane 6;

Table 2 Occurrence of OPA8₄₀₀ marker in the hemp plants analyzed

Table 2); in monoecious plants, the RAPD band was always missing, at least in all of the samples we examined. In the other cultivars and accessions tested, the association of this band with the male phenotype was constant.

Cloning of the male-specific RAPD band

The OPA8400 band was cloned, and different transformants were recovered, with inserts belonging to the three different size classes. Cloned bands of each of the three classes were used as hybridization probes on Southern blots of OPA8-amplified DNA from different male and female plants. Only one of the three classes of inserts (coded class A3, with an approximate size of 440 bp) gave as hybridization pattern a strong band at 400 bp and two minor bands, thus confirming that the clone contained the male specific fragment (Fig. 2). The amplified products from the other two insert classes did not hybridize to the OPA8 400-bp band (not shown) and were not sequenced.

RFLP analysis

Genomic DNA from 28 female and 17 male plants from five different dioecious cultivars was cut with 13 different restriction endonucleases, blotted and hybridized to $[3³²P]$ -labeled clone A3. Several different hybridization patterns were obtained; however, none of these presented any sex-linked polymorphism. An example of the patterns obtained with the enzymes *Hin*dIII, *Hin*dII, *Bgl*II and *Nco*I on 1 male and 1 female DNA of the cv 'Fibranova' is shown in Fig. 3. Polymorphism linked to the sexual phenotype was also not detected in monoecious DNA, and the hybridization signals in this case were also indistinguishable from male and female

Fig. 1a**–**c OPA8 amplification products from DNA of male (*lanes 1—5*) and female (*lanes 6—10*) individual plants of the cvs 'Carmagnola' (*panel a*), 'Fibranova' (*panel b*) and 'Uniko' (*panel c*); the *arrows* indicate the OPA8⁴⁰⁰ marker. The lane indicated with an *asterisk* shows one of the deviant phenotypes (female plant exhibiting OPA8400). *^M* Molecular weight marker (Gibco-BRL)

patterns. These results, however, indicated that the sequence corresponding to the A3 clone, derived from the male-specific RAPD band, is present in all hemp sexual phenotypes.

Fig. 3 Hybridization of the A3 clone containing the MADC2 sequence to cv 'Fibranova' DNA from one male and one female plant restricted with *Hin*dIII, *Hin*dII, *Bgl*II and *Nco*I

Sequence analysis of the male-specific marker

The cloned $OPA8₄₀₀$ marker was sequenced, and the result is shown in Fig. 4. The sequence, was named MADC2 (Male-Associated DNA from *Cannabis sativa*), according to the nomenclature proposed by Sakamoto et al. (1995), has a total length of 391 bp, possesses no open reading frames in either orientations and frames and has a $G + C$ content of 40.3%. A search in the sequence databases of GenBank, EMBL, DDBJ and PDB showed only limited homologies (ranging from 54.8% to 59.8% identity over a length of 169*—*232 bp) of MADC2 with other sequences belonging to repetitive genome regions or retrotransposon elements from different plants (barley, coconut, pine, pea, *Arabidopsis*).

Development of a SCAR marker

Two oligonucleotides were constructed from the MADC2 sequence, with sequences 5'-GTGACGTAG-

 $\overline{4}$

Fig. 4 Sequence of the MADC2 cloned fragment

GTAGAGTTGAA-3', corresponding to the positions 1–20, and 5'-GTGACGTAGGCTATGAGAG-3', corresponding to the positions 373*—*391. The effectiveness of the SCAR marker in the identification of male plants was tested on 41 plants of the cvs 'Fibranova', 'Carmagnola', 'Uniko', 'Kompolti Hybrid TC', 'Bialobrzeskie' and 'Beniko' and of one Hungarian accession. These specific primers were able to amplify a single DNA band of about 390 bp in all of the male plants tested, while in both female and monoecious plants, two bands of about 560 and 870 bp were present (Fig. 5). In the 17 staminate plants tested, the 390 bp SCAR was always amplified, while in all carpellate and monoecious plants, only the two bands of higher molecular weight were present. Out of 3 female plants showing $OPA8₄₀₀$ (Table 2) only 1 was positive when analyzed with the SCAR marker.

Discussion

The RAPD marker $OPAS₄₀₀$ appears to be adequate in screening for male plants in the hemp cultivars tested. It apparently constitutes a marker closely linked to a male-determining chromosome segment(s) of *Cannabis sativa*; as a consequence, such a marker can be fruitfully used for an early determination of sex in hemp, well before the plants screened reach the reproductive maturity. Our RAPD marker proved to be extremely reproducible under a wide variation of amplification conditions; it is clearly visible up to an annealing temperature of 44*°*C (data not shown), and changes in the origin of the primer, the *Taq* polymerase and the thermal cycler used do not affect the result. In addition, it is easily scorable in agarose gels of OPA8 amplified products from DNA minipreps (Cheung et al. 1993). Nevertheless, the reproducibility of RAPD markers has often been questioned, and currently the trend is to develop the more reliable and simple to use SCAR markers from sequenced RAPD markers (Paran and Michelmore 1993). A SCAR can indeed further simplify the massive screening of dioecious hemp cultivars and breeding lines.

Fig. 5 SCAR analysis with the male-specific primers. DNA from one male and one female plant of cvs 'Carmagnola' (*lanes 1* and *2*, respectively), 'Fibranova' (*lane 3* and *4*), 'Uniko' (*lane 5* and *6*), 'Kompolti Hybrid TC' (*lanes 7* and *8*) and 'Accession 2' (*lane 9* and *10*) was amplified with the primers constructed from the MADC2 sequence. *Lanes* 11-14 DNA from 4 different monoecious plants (cv 'Bialobrzeskie'). *M* Molecular weight markers (Gibco-BRL)

Only about 20 primers were tested before the $OPA8₄₀₀$ marker was identified. This is in agreement with Sakamoto et al. (1995), who screened only 15 decamer primers before finding two markers of 500 and 730 bp, generated by 2 different primers, that appeared to be tightly linked to the male sex. Besides, a screening of 179 RAPD primers (Operon series A*—*E and G*—*J) on male and female DNA bulks (10 individual DNAs each bulk) showed at least 12 male-associated markers (data not shown). These markers are presently being cloned and analyzed for their possible relatedness to $OPAS₄₀₀$. This situation is in sharp contrast with other studies on molecular markers linked to sex in other dioecious species. In *Pistacia vera* (Hormaza et al. 1994) about 400 primers were screened by bulked segregant analysis in order to find a 950-bp marker tightly linked to the female phenotype; in hop (*Humulus lupulus*), 1000 decamer primers were screened before the identification of a single marker useful for the construction of an efficient SCAR discriminating male and female plants (Polley et al. 1997). Finally, in asparagus (*Asparagus officinalis*), 760 primers were used in bulked segregant analysis in order to find a primer yielding two DNA fragments differentiating the two bulks, from which a SCAR marker associated with the M (male-determining) locus was developed (Jiang and Sink 1997).

The sequence corresponding to the male-specific marker is not only present in staminate plants but also in carpellate and monoecious individuals, as demonstrated by the fully comparable hybridization patterns of genomic Southern blots obtained using MADC2 as a probe. If the existence of a morphologically distinct Y chromosome is assumed, this result rules out the possibility that the sequence used to develop the marker is located solely on this chromosome. A similar result was found by Sakamoto et al. (1995); in this work, the entire length of the fragment deriving from RAPD analysis was not able to detect sex-specific polymorphisms between male and female plants, though the authors reported that a subcloned fragment deriving from the male-specific 730-bp band evidenced such polymorphism. As far as we know, these authors did not examine monoecious plants or construct a SCAR marker. The fact that, when used as an hybridization probe, the MADC2 sequence failed to discriminate the sexual phenotype might be explained by supposing that the differences at the DNA level are limited to the region including the sequence of the RAPD primer originally used; alternatively, the arrangement of the involved sequences in female plants could be such to prevent the amplification of a 400-bp fragment as a single entity.

The MADC2 sequence shares little similarity with other known sequences, except for a 50*—*60% homology with other plant sequences reported to belong to repeated regions or retrotransposon-like sequences. This limited identity is, however, interesting because it is known that DNA from sex chromosomes in different dioecious species are rich in these types of sequences (Clark et al. 1993; Grant et al. 1994). No homology was found with the MADC1 sequence (Sakamoto et al. 1995) despite a strong similarity in the $G + C$ content (39.9% for MADC1 and 40.4% for MADC2). The differences are also confirmed by the reported distribution of the restriction sites: no *Xba*I or *Xho*I sites are present in MADC2, while these enzymes were used to cut the MADC1 sequence into smaller fragments.

MADC2 is probably a non-coding genome region, as it includes no large open reading frames. On the basis of the present data, however, it is not possible to ascertain whether the sequence is part of the gene(s) for sex determination in hemp. Both the $OPA8₄₀₀$ and the SCAR marker developed from MADC2 are codominant, segregating in almost perfect accordance with sex. However, in 3 female plants out of 88 tested (Table 2) the $OPA8_{400}$ RAPD band was present, but the sex phenotype scored was female, with no apparent differences with the other $OPA8₄₀₀$ -lacking plants. This discrepancy was only partially resolved when the SCAR marker was used, as 1 of the 3 deviant female patterns was maintained; this case might be attributed to genetic recombination between the sex locus and the marker. In hemp, sex determination is complex and reported to be influenced by autosomic genes (Durand and Durand 1990) and even reversed or modified by chemical treatments or environmental factors that probably control the penetrance of the genes involved (McPhee 1924; Chailakhyan 1979; Mohan Ram and Jaiswal 1972; Mohan Ram and Sett 1979). For these reasons there is the possibility of inadequate phenotype scoring for the 3 deviant patterns observed, especially given that only female plants showed a lack of correspondence between presence/absence of $OPAS₄₀₀$ and sex.

The described RAPD and SCAR markers proved to be extremely effective in the discrimination of male plants of *Cannabis sativa*. In fact, in hemp the only chance of early sex discrimination has up to now relied on complex multivariate analysis performed on vegetative characters (Lacombe 1980). These markers might be fruitfully used for an early screening of plants to be crossed in a breeding program, in combination with the scoring of other agronomically important characters typical of this species, such as the low cannabinoid content and the high content and/or quality of the fibres, and will, therefore, be of great potential utility in the marker-assisted selection in this crop.

Acknowledgments This work was supported by the Italian Ministry of Agriculture, in the framework of the project ''Induction of phenotypic markers and improvement of common hemp''.

References

- Abel E (1990) Marihuana: the first 12,000 years. Plenum Press, New York
- Allavena D (1961) Fibranova, nuova varietà di canapa ad alto contenuto di fibra. Sementi Elette 7 : 34*—*44
- Altschul SF, Gish W, Miller E, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215 : 403*—*410
- Bócsa I (1994) Interview. In: Journal of the International Hemp Association, vol.1, no.2, pp 61*—*63
- Bredemann G (1938) Züchtung des Hanfes auf Fasergehaltes. Die Ergebnisse des Jahres 1937. Faserforschung 4 : 239*—*258
- Chailakhyan MK (1979) Genetic and hormonal regulation of growth, flowering, and sex expression in plants. Amer J Bot 66 : 717*—*736
- Cheung WY, Hubert N, Landry BS (1993) A simple and rapid DNA microextraction method for plant, animal and insect suitable for RAPD and other PCR analysis. PCR Methods Applic 3 : 69*—*70
- Clark MS, Parker JS, Ainsworth CA (1993) Repeated DNA and heterochromatine structure in *Rumex acetosa*. Heredity 70 : 527*—*536
- Durand R, Durand B (1990) Sexual determination and sexual differentiation. Crit Rev Plant Sci 9 : 295*—*316
- Faeti V, Mandolino G, Ranalli P (1996) Genetic diversity of *Cannabis sativa* L. germplasm based on RAPD markers. Plant Breed 115 : 367*—*370
- Feinberg AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activities. Anal Biochem 132 : 6*—*13
- Grant S, Houben A, Vyskot B, Siroky J, Pan WH, Macas J, Saedler H (1994) Genetics of sex determination in flowering plants. Dev Genet 15 : 214*—*230
- Hormaza JI, Dollo L, Polito VS (1994) Identification of a RAPD marker linked to sex determination in *Pistacia vera* using bulk segregant analysis. Theor Appl Genet 89 : 9*—*13
- Jiang C, Sink KC (1997) RAPD and SCAR markers linked to the sex expression locus M in asparagus. Euphytica 94 : 329*—*333
- Lacombe JP (1980) Discrimination des sexes en fonction de caractères végétatifs précoces chez la Chanvre dioïque (*Cannabis sativa* L.). Physiol Veg 18 : 419*—*430
- McPhee HC (1924) The influence of environment on sex in hemp, *Cannabis sativa* L. J Agric Res 28 : 1067*—*1080
- Mohan Ram HY, Jaiswal VS (1972) Induction of male flowers on female plants of *Cannabis sativa* by gibberellins and its inhibition by abscisic acid. Planta 105 : 263*—*266
- Mohan Ram HY, Sett R (1979) Sex reversal in the female plants of *Cannabis sativa* by cobalt ions. Proc Indian Acad Sci 88 : 303
- Mohan Ram HY, Sett R (1985) *Cannabis sativa*. In: Halevy handbook of flowering, vol. 2. CRC Press, Boca Raton, Fla., pp 131
- Paran I, Michelmore RW (1993) Development of PCR-based markers linked to downy mildew resistance in lettuce. Theor Appl Genet 85 : 985*—*993
- Polley A, Seigner E, Ganal MW (1997) Identification of sex in hop (*Humulus lupulus*) using molecular markers. Genome 40 : 357*—*361
- Sakamoto K, Shimomura K, Komeda H, Satoh S (1995) A male associated DNA sequence in a dioecious plant, *Cannabis sativa* L. Plant Cell Physiol 36 : 1549*—*1554
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, vol 2, 2nd edn. Cold Spring Harbor Laboratory Press, New York, pp 9 : 31*—*9 : 34
- Taylor B, Powell A (1985) Isolation of plant DNA and RNA. Focus 4:4*—*6
- Yamada I (1943) The sex chromosome of *Cannabis sativa* L. Seiken Ziho 2 : 64*—*68